Research Article

Phytochemical Analysis

Received: 5 August 2008, Revised: 14 March 2009, Accepted: 26 March 2009 Published online in Wiley Interscience: 16 June 2009

(www.interscience.wiley.com) DOI 10.1002/pca.1137

# A Retention Index Calculator Simplifies Identification of Plant Volatile Organic Compounds

**Mary Lucero,[a]* Rick Estell,[a] María Tellez[b] and Ed Fredrickson[a]**

**ABSTRACT:**
**Introduction – Plant volatiles (PVOCs) are important targets for studies in natural products, chemotaxonomy and biochemical ecology. The complexity of PVOC profiles often limits research to studies targeting only easily identified compounds. With the availability of mass spectral libraries and recent growth of retention index (RI) libraries, PVOC identification can be achieved using only gas chromatography coupled to mass spectrometry (GCMS). However, RI library searching is not typically automated, and until recently, RI libraries were both limited in scope and costly to obtain.**
**Objective – To automate RI calculation and lookup functions commonly utilised in PVOC analysis.**
**Methodology – Formulae required for calculating retention indices from retention time data were placed in a spreadsheet along with lookup functions and a retention index library. Retention times obtained from GCMS analysis of alkane standards and *Koeberlinia spinosa* essential oil were entered into the spreadsheet to determine retention indices. Indices were used in combination with mass spectral analysis to identify compounds contained in *Koeberlinia spinosa* essential oil.**
**Results – Eighteen compounds were positively identified. Total oil yield was low, with only 5 ppm in purple berries. The most abundant compounds were octen-3-ol and methyl salicylate. The spreadsheet accurately calculated RIs of the detected compounds.**
**Conclusion – The downloadable spreadsheet tool developed for this study provides a calculator and RI library that works in conjuction with GCMS or other analytical techniques to identify PVOCs in plant extracts. Copyright © 2009 John Wiley & Sons, Ltd.**

**Keywords:** Kováts index; gas chromatography; *Koberlinia spinosa*; essential oil

## Introduction

Plant volatile organic compounds (PVOCs) are commonly identified using gas chromatography (GC) in combination with (1) mass spectroscopy (MS) for spectral matching, (2) flame ionisation detection (FID) for quantification of peak areas from diverse unknowns and (3) retention index (RI) matching. A single plant extract may contain hundreds of structurally similar terpenoids that produce highly similar mass spectra. Thus, compound matching by mass spectrometry alone is error prone. Matching by comparing retention times with those of standards run on the same system can be impractical, since many plant compounds are not commercially available. Published retention indices are more suitable than retention times for comparison across different chromatographic systems. The combined matching of mass spectra and retention indices can facilitate positive identification of more than 100 compounds from a single plant extract with only a few chromatograms (Tellez *et al*., 1997a).

While mass spectral library search capabilities have been standard components of software sold with GCMS systems for decades, standard libraries of retention indices have only recently been added to major chemical databases (NIST Mass Spec Data Center, 2005a, b; Babushok *et al*., 2007). Hence, retention index library matching has not traditionally been included as a feature in major chromatography software systems.

For several years, our research group has examined PVOC profiles in conjunction with livestock herbivory and general arid land ecology (Fig. 1). Throughout these studies, we have relied heavily on Robert Adams's mass spectral libraries, which include both mass spectra and retention indexes (Adams, 2001, 2007). To automate retention index calculation and matching in conjuction with this library, we developed a simple spreadsheet tool which uses functions contained within Microsoft Excel® and OpenOffice Calc (OpenOffice.org™ v. 3.0.1) to calculate retention indices of unknowns by comparing them with retention times of hydrocarbon alkanes analysed with the same separation method. After calculating retention indices, the tool searches input libraries, returning compounds with retention indices near those of the peak in question. An input library consisting of retention indices reported for PVOC compounds examined in the PVOC analyses referenced in Fig. 1 is included. Additional compounds may be added to this library by the end user.

Although similar tools are increasingly likely to be integrated within modern chromatography software systems, we believe this tool will be of use to research, quality control and academic laboratories that lack access to the latest chromatographic software and systems.

* Correspondence to: M. Lucero, USDA, ARS, Jornada Experimental Range, 2995 Knox Street Las Cruces, NM 88003, USA. E-mail malucero@nmsu.edu

[a] USDA, ARS, Jornada Experimental Range, 2995 Knox Street Las Cruces, NM 88003, USA

[b] Inter Science Institute, 944 West Hyde Park Boulevard, Inglewood, CA 90302, USA

Here we demonstrate the utility of our RI Calculator tool by determining retention indices of volatile compounds associated with the essential oil of *Koeberlinia spinosa* Zucc. ('crucifixion thorn', KOSP), a thorny, nearly leafless arid land species noted for upward growing, putatively hydrotropic roots (Gibbens and Lenz, 2001). In doing so, we demonstrate the utility of the RI Calculator for identifying volatiles from complex profiles for diverse applications.

## Experimental

### Plant sampling, oil extraction and instrumental analysis

KOSP samples were collected from the JER near Las Cruces, New Mexico at approximately N 32°36.607 latitude and W 106°33.384 longitude, 1726 m above sea level (Fig. 1). The current year's growth (approximately 1 Kg fresh weight) was removed from randomly selected branches of 10 individual plants and combined to form composite samples. Composited samples were cut by hand to lengths not exceeding 2.54 cm, and triplicate 20 g aliquots of this material were extracted by steam distillation as previously described (Lucero *et al*., 2005). Briefly, chopped green tissue was placed in a round-bottom flask containing 100 mL of distilled water and attached to a Likens–Nickerson distillation apparatus, where it was distilled for 6 h. Pentane (10 mL) was added to the U-tube and to a pear-shaped collection flask. The pentane was maintained at 60–70°C throughout the 6 h distillation. Following distillation an additional 8 mL of pentane was used to rinse the apparatus. Pentane fractions were combined, dried over anhydrous magnesium sulphate, and filtered. The solvent was removed using a rotary evaporator under reduced pressure.

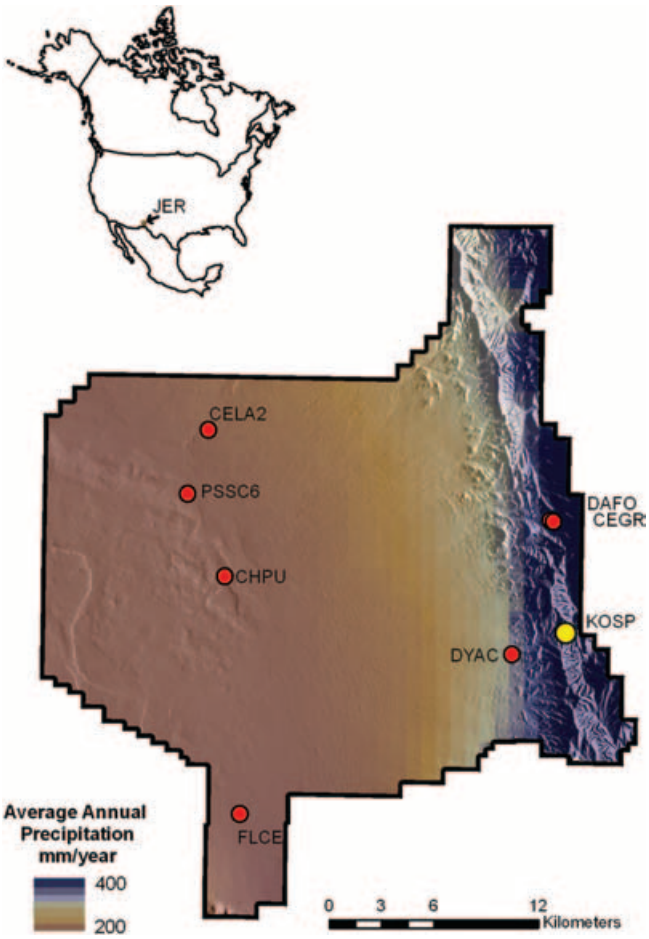



**Figure 1.** Volatile compounds extracted from shrubs growing on the USDA-ARS Jornada Experimental Range, located in the southwestern USA, have provided a library of retention indices against which newly analysed volatiles can be compared. Sample sites for each species are indicated with the NRCS species codes as follows: *Ceanothus greggii* Gray (CEGR) (Lucero *et. al*., 2009); *Ceratoides lanata* (Pursh) J. T. Howell var. *subspinosa* (Rydb.) J. T. Howell (CELAS2) (Lucero *et al*., 2004); *Chrysothamnus pulchellus* (A. Gray) Greene (CHPU4) (Tellez *et al*., 1998); *Dalea formosa* Torr. (DAFO) (Lucero *et al*., 2005); *Dyssodia acerosa* DC. (DYAC) (Tellez *et al*., 1997a); *Koberlinia spinosa* (KOSP), *Flourensia cernua* (FLCE) (Tellez *et al*., 1997b) and *Psorothamnus scoparius* (A. Gray) Rydb. (PSSC6) (Lucero *et al*., 2003).

Fruit was harvested from the same plants at both immature (green) and mature (purple) stages. Whole berries were steam distilled in the same manner, except that only 10 g of tissue were used in each replicate. Oil retrieved from each distillation was diluted in 100% ethanol for GCMS analysis, or injected as pure oil for GC-FID analysis.

Solid-phase microextraction (SPME) analyses was conducted by placing 0.5 g of finely chopped tissue into 4 mL screw-top vials sealed with poly(tetrafluoroethylene) (PTFE)/silicon septa (Supelco®). The vials were equilibrated at 30°C for 2 h and then exposed to 1 cm of a 100 µm PDMS fiber (Supelco®), for 10 min. The fibre was immediately injected into the appropriate gas chromatograph inlet to a depth of 3 cm. The fibre remained in the injector for 5 min to remove residual volatiles, and blank runs were performed after each sample.

GCMS analysis was performed using a Varian model 3400 GC with a DB-5 column (30 m × 0.25 mm fused silica capillary column, film thickness 0.25 µm) coupled to a Finnigan ion trap mass spectrometer (EI, 70 eV). Helium (1 mL/min) was used as a carrier gas, and injector and transfer line temperatures were set at 220 and 260°C, respectively. The initial column temperature was 60°C, and a linear temperature increase of 3°C/min was programmed into each 65 min run. When injecting oils, a series of large (500 ng) to small (50 ng) injections were used to validate retention times for both low- and high-concentration components.

To validate peak area percentages revealed on the total ion chromatogram, the oil was also analysed using a Shimadzu GC8APF equipped with a flame ionisation detector and fitted for use with capillary columns. A split/splitless injector was used, and the column type, temperature gradients, and helium flow rates were identical to those used for the GCMS analysis. The injector temperature was 250°C.

Dry matter percent was determined by drying triplicate, 2 g samples of chopped tissue at 60°C for 24 h. Oil yield was reported as µg oil/gram dry tissue.

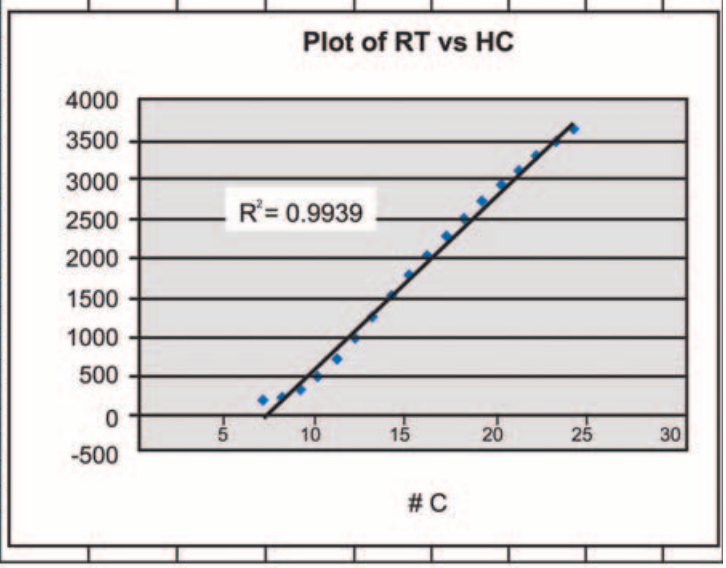
### Retention time and retention index matching using the 'RI Calculator'

To facilitate repeated calculation of retention indices, a standard containing 5 ng/µL each of 18 *n*-alkane hydrocarbons (heptane through *n*-pentacosane) was run prior to analysis of essential oil extracts using the instrument parameters described above. The retention times at which each hydrocarbon alkane standard eluted was recorded on the Hydrocarbon RTs page of the 'RI Calculator' spreadsheet (Fig. 2). A read-only copy of the



| Name of Hydrocarbon | Number of Carbon Atoms | Seconds | Current RT in minutes | Typical RT in seconds | Typical RT in minutes |
|---|---|---|---|---|---|
| heptane | 7 | 105.78 | 1.763 | | |
| octane | 8 | 157.5 | 2.625 | 167 | 2.78 |
| nonane | 9 | 255.78 | 4.263 | 263 | 4.38 |
| decane | 10 | 423.36 | 7.056 | 422 | 7.03 |
| undecane | 11 | 643.92 | 10.732 | 635 | 10.58 |
| dodecane | 12 | 903.42 | 15.057 | 883 | 14.72 |
| n-tridecane | 13 | 1177.14 | 19.619 | 1143 | 19.05 |
| n-tetradecane | 14 | 1448.94 | 24.149 | 1401 | 23.35 |
| n-pentadecane | 15 | 1711.44 | 28.524 | 1651 | 27.52 |
| n-hexadecane | 16 | 1961.88 | 32.698 | 1891 | 31.52 |
| n-heptadecane | 17 | 2200.02 | 36.667 | 2120 | 35.33 |
| n-octadecane | 18 | 2426.52 | 40.442 | 2338 | 38.97 |
| n-nonadecane | 19 | 2641.62 | 44.027 | 2547 | 42.45 |
| icosane | 20 | 2846.82 | 47.447 | 2746 | 45.77 |
| henicosane | 21 | 3042.42 | 50.707 | 2937 | 48.95 |
| n-Docosane | 22 | 3229.2 | 53.82 | 3120 | 52.00 |
| n-tricosane | 23 | 3408.18 | 56.803 | 3295 | 54.92 |
| n-tetracosane | 24 | 3579.48 | 59.658 | 3464 | 57.73 |
| n-pentacosane | 25 | 3747.72 | 62.462 | 3627 | 60.45 |




1) **Enter the retention times obtained on your GC system for the alkane standards listed above.** Retention times recorded in minutes can be entered into the 'Current RT in minutes' column. Retention times in seconds will be calculated automatically and used downstream to calculate retention indices.
1a) If preferred, retention time in seconds can be placed directly on the seconds column, overwriting the 'zeros' that appear when the minutes column is blank.
1b) If values are entered directly as seconds, the 'minutes' column can be left blank.

2) As the complete series of alkane retention times is entered, **verify that the line developing on the graph has an $R^2$ value > 0.99.** Lower $R^2$ values may indicate an insufficent number of standards or problems with chromatography and peak identification.
2a) Heptane may not be detected due to solvent delays or may be hidden in the solvent peak. Although it is included in the series, a heptane retention time can easily be omitted without a significant drop in the $R^2$ value of the plotted line.

**Figure 2.** The HYDROCARBON RTs page of the RI Calculator workbook. Retention time values from hydrocarbon standards can be pasted into either column C (seconds) or column D (minutes). New values can quickly be compared with retention times obtained from previous runs, so that deviations can be identified. The $R^2$ value of the line associated with the scatter plot at the bottom of the page will illustrate the linearity of the retention time vs hydrocarbon number plot, which should exceed 0.99. This figure is available in colour online at www.interscience.wiley.com/journal/pca

spreadsheet (supplemental data) can be opened in recent versions of Microsoft Excel® or Open Office Calc™.

Mean *n*-alkane retention times obtained by separating standards as described above are included in the spreadsheet as reference points. This separation method allows the user to compare results to retention indices reported in Robert Adams's database of essential oil retention indices (Adams, 2001), or to retention indices determined by our group and included in the calculator. A scatter plot of hydrocarbon numbers against retention times appears on a graph within the spreadsheet, along with a linear plot and the $R^2$ value of the linear regression. This graph helps to assess linearity of the *n*-alkane retention times. Outlier points or $R^2$ values significantly less than 1 may indicate problems with the chromatographic system or with compound identification. Our group routinely achieves $R^2$ values greater than 0.99 with the standards and methods described herein.

Once the *n*-alkane retention times and linearity relationships are properly determined, the next step is to transfer the retention times of unknown compounds, separated with the same instrument parameters from the GC output to the appropriate column (A for seconds, B for minutes) on the 'sample data entry' page (Fig. 3). If GC retention times are reported in minutes, they can be pasted into the 'minutes' column, and the seconds will be calculated automatically. If the retention time is provided in seconds, values can be pasted directly into the 'seconds' column, overwriting the minute conversion formulae contained therein. In this case, the minute column is ignored, since it is not needed for subsequent calculations. The column void time must be manually entered in the upper right-hand corner.

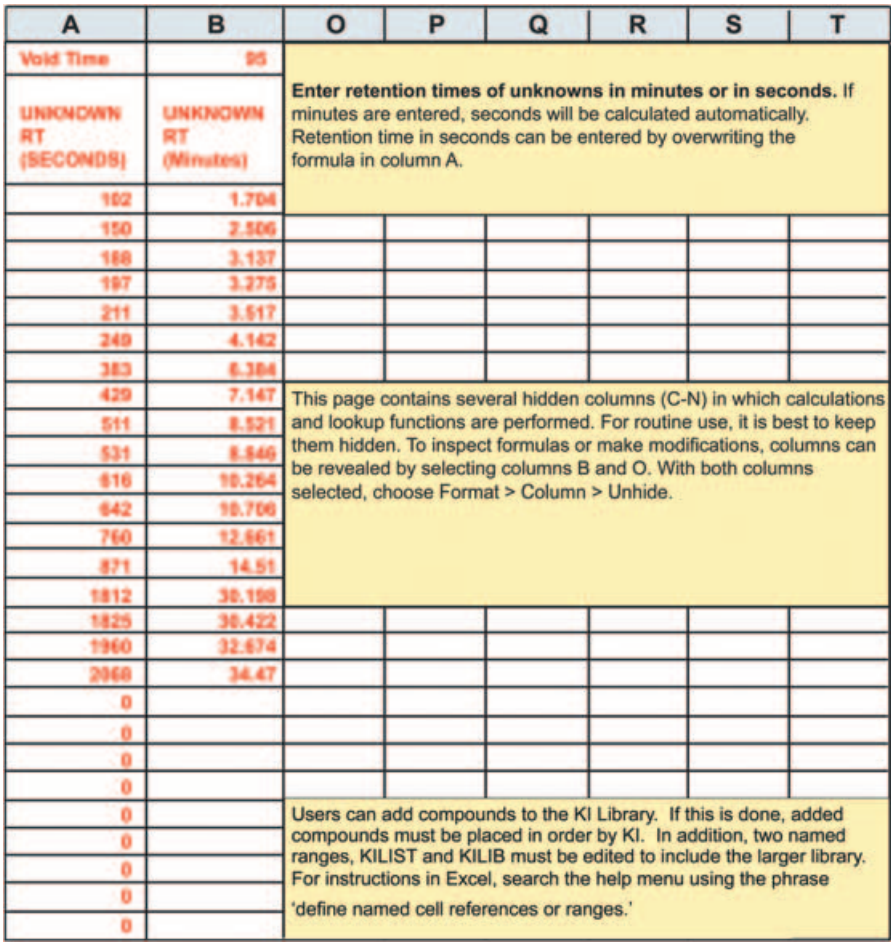
Remaining spreadsheet columns were hidden to simplify viewing. The hidden columns contain lookup, matching and index formulae used to identify retention times of the nearest alkanes (taken from the Hydrocarbon RTs page), and the formula (A) used to calculate retention indices. In order to review these formulas, a user can reveal hidden columns simply by selecting Format > Columns > Unhide from the top menu.

Formula A:

$$RI = \frac{100 * n + [\log(RT_{unknown} - v) - \log(RT_{smaller\,alkane} - v)]}{\log(RT_{larger\,alkane} - v) - \log(RT_{smaller\,alkane} - v)}$$

where $n$ = the number of C in the smaller alkane; $RT_{unknown}$ = the retention time of the unknown in seconds; $v$ = column void time; $RT_{smaller\ alkane}$ = the retention time of the smaller alkane; and $RT_{larger\ alkane}$ = the retention time of the larger alkane.

The results are viewed by selecting the 'Results' tab from the bottom of the workbook. The output shows the RI calculated for



| A | B | O | P | Q | R | S | T |
|---|---|---|---|---|---|---|---|
| Void Time | 95 | | | | | | |
| UNKNOWN RT (SECONDS) | UNKNOWN RT (Minutes) | | | | | | |
| 102 | 1.704 | | | | | | |
| 150 | 2.506 | | | | | | |
| 188 | 3.137 | | | | | | |
| 197 | 3.275 | | | | | | |
| 211 | 3.517 | | | | | | |
| 249 | 4.142 | | | | | | |
| 383 | 6.384 | | | | | | |
| 429 | 7.147 | | | | | | |
| 511 | 8.521 | | | | | | |
| 531 | 8.846 | | | | | | |
| 616 | 10.264 | | | | | | |
| 642 | 10.706 | | | | | | |
| 760 | 12.661 | | | | | | |
| 871 | 14.51 | | | | | | |
| 1812 | 30.198 | | | | | | |
| 1825 | 30.422 | | | | | | |
| 1960 | 32.674 | | | | | | |
| 2068 | 34.47 | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |
| 0 | | | | | | | |



Figure 3. The Sample Data Entry page of the RI Calculator. Columns C–N, used strictly for calculation of values derived from the entries, are hidden. They can be viewed by selecting Format > Columns > Unhide from the top menu. This figure is available in colour online at www.interscience.wiley.com/journal/pca

each unknown retention time entered. In addition, the columns to the right show the compound obtained from the page 'RI Library' with a retention time nearest to the unknown. The published reference in which the RI for the matching compound was determined is shown, along with the USDA-NRCS Plants Database symbol for the plant species in which the library compound was identified (Fig. 4).

## Compound identification

Positive identification of organic compounds generally requires application of two or more analytical techniques. In PVOC analysis, combining mass spectral and retention index matching is convenient since both can be carried out with a single chromatographic run. In the absence of spectral analysis, or when several compounds have similar spectra and retention indices on a given column, matching of retention indices determined on columns with different stationary phase polarities may suffice. In the current experiment, mass spectral matching was obtained using Magnum™ software. A positive match required both a spectral fit score of ≥950 and a retention index within five units of reported values.

Some judgement on the part of the analyst is necessary for determining acceptable variation between reported and observed retention indices. For example, compounds that are poorly retained (RIs < 850) on the column are likely to show more variable retention times than those with indices between 850 and 1900 using the DB-5 column described above. Late eluting compounds (RIs > 1900) may undergo band broadening which may cause variation in assessment of retention times. Thus, a 5 unit difference between reported and observed retention indices would be more acceptable for very early or very late eluting compounds than for compounds with intermediate retention times.

A second factor to consider when evaluating differences between reported and observed retention times is the difference between reported and observed retention times for other compounds within the same extract. For example, if an extract contains tricyclene (RI = 930), $\alpha$-pinene (RI = 942), and an unknown compound with a spectra matching camphene and a retention index of 950, the unknown is not likely to be camphene. While the calculated retention index is within 2 units of the 953 index reported for camphene in *Dalea formosa* extracts (Lucero *et al.*, 2005), the indices observed for tricyclene and $\alpha$-pinene in the same extract are higher than the mean retention indices reported for those compounds in other literature (Tellez *et al.*, 1997b; Adams, 2001; Medina *et al.*, 2005); hence, one would expect camphene, which elutes within a minute of tricyclene and $\alpha$-pinene, to also have a slightly higher-than-average index. The various retention indices reported for compounds identified in various plant extracts



| A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|
| UNKNOWN RT (Minutes) | UNKNOWN RT (Seconds) | UNKNOWN KI | Nearest KI from Local Library | Library compound with KI nearest that of the unknown | Source Plant Species Code | Reference | | | |
| 1.70 | 102 | #N/A | #N/A | #N/A | #N/A | #N/A | | | |
| 2.51 | 150 | 793 | #N/A | #N/A | #N/A | #N/A | | | |
| 3.14 | 188 | 842 | 835 | isovaleric acid | CEGR | Lucero *et al.*, Journal of essential oil research (under review) | | | |
| 3.28 | 197 | 851 | 853 | (E)-2-hexenal | ANCA | Medina *et al.*, Journal of Agricultural and Food Chemistry 53 (2005) 8694. | | | |
| 3.52 | 211 | 865 | 868 | n-hexanol | CEGR | Lucero *et al.*, Journal of essential oil research (under review) | | | |
| 4.14 | 249 | 895 | 899 | n-heptanal | CEGR | Lucero *et al.*, Journal of essential oil research (under review) | | | |
| 6.38 | 383 | 982 | 982 | -pinene | DAFO | Lucero *et al.*, Journal of essential oil research 17 (2005) 645. | | | |
| 7.15 | 429 | 1003 | 1004 | a-phellandrene | CHPU | Tellez *et al.*, Journal of Essential Oils Research 10, 201 (Mar/Apr 1998). | | | |
| 8.52 | 511 | 1046 | 1045 | phenylacetaldehyde | PSSC | Lucero *et al.*, Journal of Essential oil Research 15 (2003) 108. | | | |
| 8.85 | 531 | 1055 | 1052 | (E)-ocimene | ANCA | Medina *et al.*, Journal of Agricultural and Food Chemistry 53 (2005) 8694. | | | |
| 10.26 | 616 | 1090 | 1089 | p-mentha-2,4(8)diene | CEGR | Lucero *et al.*, Journal of essential oil research (under review) | | | |
| 10.71 | 642 | 1099 | 1099 | linalool | ANCA | Medina *et al.*, Journal of Agricultural and Food Chemistry 53 (2005) 8694. | | | |
| 12.66 | 760 | 1149 | 1147 | trans-verbenol | CEGR | Lucero *et al.*, Journal of essential oil research (under review) | | | |
| 14.51 | 871 | 1189 | 1189 | a-terpineol | DYAC | Tellez *et al.*, Journal of Agricultural and Food Chemistry 45, 3276 (1997). | | | |
| 30.20 | 1812 | 1542 | 1538 | a-cadinene | CHPU | Tellez *et al.*, Journal of Essential Oils Research 10, 201 (Mar/Apr 1998). | | | |
| 30.42 | 1825 | 1547 | 1548 | elemol | FLCE | Tellez *et al.*, Journal of essential oil research 9, 619 (Nov/Dec 1997). | | | |
| 32.67 | 1960 | 1599 | 1607 | b-oplopenone | FLCE | Tellez *et al.*, Journal of essential oil research 9, 619 (Nov/Dec 1997). | | | |
| 34.47 | 2068 | 1646 | 1646 | a-muurolol | DYAC | Tellez *et al.*, Journal of Agricultural and Food Chemistry 45, 3276 (1997). | | | |

Plant species symbols match the symbols used by the USDA-NRCS PLANTS database to abbreviate names of species in which the compound was detected.

RI's in the library were determined from compounds detected within the indicated species. References to the published articles describing the analysis and identification of the matching library compounds is provided.

Figure 4. The Results page of the RI Calculator. Unknowns, listed by retention time, are shown with calculated RIs. In addition, the library compound with the nearest RI is provided, along with the USDA-NRCS species code of the source plant in which the library compound was previously detected. References describing identification of the library compound are shown. This figure is available in colour online at www.interscience.wiley.com/journal/pca

identified on the JER can be seen on the 'RI Library' page of the RI Calculator (see Supporting Information).

### Quantitative analysis

Relative peak areas eluting from FID chromatograms with retention indices matching those of compounds identified in parallel GCMS chromatograms were used to estimate micrograms of compound per gram dry plant tissue (ppm).

## Results and Discussion

### Essential oils of *Koeberlinia spinosa*

KOSP is the only species within the genus *Koeberlinia*. The nearly leafless shrub produces photosynthetic stems and sharp thorns which effectively deter large mammals. However, brightly colored berries which appear in the summer are utilised by small herbivores, including birds and insects. The absence of leaves, which are the source of many PVOCs in most plants, suggests low overall volatile production.

Triplicate distillations of the photosynthetic, leafless stem tissue yielded less than 0.1 ppm (μg oil/g dry matter). Yields for the green and purple berries were 1 and 5 ppm, respectively. Low yield from stem tissue prevented direct analysis of oil samples. Therefore, SPME headspace analysis was used to determine the volatile composition. SPME generally provides greater sensitivity for qualitative analysis than steam distillation, but can be more difficult to analyse quantitatively (Lucero *et al*., 2006). For this reason, the amounts of volatiles detected in stem extracts are reported simply as peak area percentage (Table 1).

The RI Calculator in combination with mass spectral and retention index libraries facilitated identification of 15 compounds in green berries, six compounds in purple berries and two compounds in the stem essential oil from *Koeberlinia spinosa* essential oil. Seven of these compounds have not been detected in other species collected from the JER (Table 1). Thirty-seven additional compounds were reproducibly detected, but could not be positively identified (data not shown). Most of these were small peaks with low signal-to-noise ratios, thus poor spectral purity prevented matching to library compounds. However, berries also contained several large, spectrally impure peaks with retention indices of 2568, 3239, 3301, 3404 and 3740. These five peaks comprised 20 and 35% of the total FID peak areas for green and purple berries, respectively, but impurities prevented spectral matching and positive identification.

Not surprisingly, oil composition differed by tissue and by maturity. Only two clearly identifiable compounds, *n*-octanol and *n*-decanal, were observed in stem tissues. Both of these compounds were also present in the green berries, which provided the most complex mixture of detectable volatiles. Three compounds, methylbenzoate, methyl salicylate and *β*-eudesmol were present in both green and purple berries. All three of these were more abundant in the green berries. Compounds identified only in the purple berries included 6-methyl-5-hepten-2-one, 8-*α*-acetoxyelemol and hexadecanol.



**Table 1.** *Koeberlinia spinosa* volatiles identified by RI and spectral matching. Asterisks identify compounds that have not previously been observed in plant species from the Jornada Experimental Range

| Identified compounds | RI | ppm in green berries | ppm in mature berries | Percentage area stem |
|---|---|---|---|---|
| Octen-3-ol* | 981 | 0.537 | | |
| 6-Methyl-5-hepten-2-one | 983 | | 0.047 | |
| 6-Methyl-5-hepten-2-ol* | 994 | 0.018 | | |
| *p*-Methylanisole* | 1021 | <0.001 | | |
| *n*-Octanol* | 1072 | 0.010 | | 0.89 |
| Methybenzoate | 1091 | 0.068 | 0.012 | |
| Methyl salicylate | 1191 | 0.114 | 0.011 | |
| *n*-Decanal | 1204 | 0.021 | | 0.47 |
| (*E*)-2-decenal | 1262 | 0.014 | | |
| 2*E*,4*E*-decadienal* | 1314 | 0.026 | | |
| Geranyl acetone* | 1453 | 0.013 | | |
| Elemol | 1549 | 0.042 | | |
| Eremoligensol* | 1629 | 0.024 | | |
| Muurolol | 1641 | 0.024 | | |
| $\beta$-Eudesmol | 1651 | 0.055 | 0.013 | |
| 8-$\alpha$-Acetoxyelemol* | 1789 | | 0.009 | |
| Hexadecanol* | 1879 | | 0.034 | |
| Methyl hexadecanoate* | 1927 | 0.041 | | |

### Utility of the RI Calculator

The RI Calculator spreadsheet effectively simplified the task of determining retention indices for unknown compounds. The lookup feature, which compares calculated RIs with those previously reported, was useful for narrowing the list of possible spectral matches to those that also had matching RIs. However, because the lookup feature only identifies the nearest library compound with an RI lower than the reported RI, manually scrolling through the library for similar RIs and checking outside references for plant volatile retention indices was still necessary.

The ability to identify volatile compounds in complex GC chromatograms with nothing more than spectral library and retention index matching is neither new nor novel, but is arguably underutilised in chemotaxonomy and other areas of chemical ecology (Adams, 2001). Retention indices, introduced 50 years ago, are robust constants that convey distinct information about the compounds they describe (Kováts, 1958). Yet, unlike chemical mass spectra, for which reference libraries abound, retention index databases have been slow to catch on. Robert Adams's mass spectral library of plant compounds is one of only a few that included retention indices with its early editions (Adams, 2001).

It is possible that hesitancy to incorporate retention indices with databases serving multiple users arose due to the rapid changes in chromatography during that time. When Kováts first proposed use of retention indices, his research was based on compounds separated isothermally on packed GC columns. However, more powerful temperature gradient separations and fused silica capillary columns were rapidly becoming the industry standard. Hence, modifications to retention index determinations had to be incorporated (Kováts, 1965). In the decades that followed, numerous new stationary phases became available for chemical separations. Because phase material strongly influences the retention index of a compound, the rapid development of novel phase technologies may have temporarily discouraged efforts to invest in development and publication of retention index libraries, which would only be of value to end users who chose the same phase. However, the added value of retention properties for compound identification is indisputable (Ettre, 2003; Harangi, 2003). Retention index data for numerous PVOCs have been available in commercially available libraries for many years (Adams, 2001). More recently, retention index data have been added to readily accessible public databases such as the NIST Chemistry WebBook (Babushok *et al*., 2007).

The retention index library offered in the RI Calculator (Supporting Information) can be readily expanded to include additional compounds of interest to the end user. To expand the library, simply add new compounds with reported retention indices to the list on the workbook sheet entitled 'RI Library'. The entire list should then be selected and re-sorted so that all compounds are listed in order by retention index, followed by compound name. Finally, to ensure that lookup commands reference the expanded list of entries, choose Insert > Name > Define from the menu on the top of the page, and select the name 'RILIST' from the 'Define Name' box that appears. While RILIST is selected, drop to the 'refers to' box and expand the range of cells in column A to include the new entries. Next, select the name 'RILIB' and repeat this process, selecting a new range of cells from columns A–D to include the new library entries. Now, all added compounds will be searchable.

The RI Calculator presented here evolved over the course of several studies of essential oil compositions. Users interested in volatiles with lower molecular weights will need to replace the hydrocarbon standards listed on the Hydrocarbon RTs page of the calculator with appropriate standards having molecular weights above and below the molecular weights of their compounds of interest. Such compounds are not represented in the existing RI library, but could be added by the end user as described above. The lookup function for matching calculated retention indices to those previously reported is limited, in that it only returns the lowest-value compound from the library list that meets the 'closest hit' criteria. A solution to this weakness could be obtained by building the tool in a database, rather



than on a spreadsheet. For example, in Microsoft Access®, all matching compounds could be displayed as a drop down list, and the best match could be selected. However, we find a greater number of users familiar with spreadsheets than with databases. Hence, we felt the spreadsheet-based application would be easier for individuals to modify and adapt to their unique laboratory needs. The utility of the spreadsheet in OpenOffice Calc is an added benefit, since OpenOffice Calc is open source software freely available to the general public. Because the authors utilised the spreadsheet in Microsoft Excel®, it has been less thoroughly tested in OpenOffice Calc. Spreadsheet copies available as Supporting Information are saved in a read-only format. Users' data that is entered can be saved by renaming the file using the 'Save As' function. This prevents overwriting the original formats and formula.

### Supporting information

Supporting information can be found in the online version of this article.

### Acknowledgements

The authors thank Helena Deswood and Yuan Feng Wang for assistance with *Koberlinia spinosa* essential oil analysis, and Barbara Nolen for map construction. We thank Ramona Parra and Conrad Fjetland for helpful reviews of the initial draft, and Tom Popham for assistance with spreadsheet lookup functions. The Retention Index calculator was constructed in Microsoft Excel®, and has also been tested in OpenOffice Calc (OpenOffice.org v. 3.0.1). Other spreadsheet software is expected to be equally suitable for this application. Use of the specific software packages described herein does not constitute official endorsement or approval by the United States Department of Agriculture or the National Agricultural Library of any product or service to the exclusion of others that may be suitable.

## References

Adams RP. 2001. *Identification of Essential Oil Components by Gas Chromatography/quadrupole Mass Spectroscopy*. Allured: Carol Stream, IL.

Adams RP, ed. 2007. *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry*, 4th edn. Allured: Carol Stream, IL.

Babushok VI, Linstrom PJ, Reed JJ, Zenkevich IG, Brown RL, Mallard WG and Stein SE. 2007. Development of a database of gas chromatographic retention properties of organic compounds. *J Chromatography A* **1157**: 414–421.

Ettre LS. 2003. Retention Index Expressions. *Chromatographia* **58**: 491–494.

Gibbens RP and Lenz JM. 2001. Root systems of some Chihuahuan Desert plants. *J Arid Environ* **49**: 221–263.

Harangi J. 2003. Retention index calculation without *n*-alkanes—the virtual carbon number. *J Chromatography A* **993**: 187–195.

Kováts E. 1958. Gas-chromatographische charakterisierung organischer Verbindugen Teil 1: Retentionsindices aliphatischer halogenide, alkohole, aldehyde und ketone. *Helvetica Chimica Acta* **41**: 1915–1932.

Kováts E. 1965. Gas chromatographic characterization of organic substances in the retention index system. In *Advances in Chromatography*, Giddings JC, Keller RA (eds). Marcel Dekker: New York; 229–247.

Lucero ME, Estell RE and Fredrickson EL. 2003. The essential oil composition of *Psorothamnus scoparius* (A. Gray) Rydb. *J Essen Oil Res* **15**: 108–111.

Lucero M, Estell R, Anderson D, Fredrickson E and Remmenga M. 2004. *Differences in Volatile Profiles Between Populations if Ceratoides Lanata Var. Subspinosa (Rydb.) J.T. Howell*, Sosebee RE, Wester D, Britton C, McArthur ED (eds). US Department of Agriculture, Forest Service, Rocky Mountain Research Station: Lubbock, TX; 142–146.

Lucero ME, Estell RE and Sedillo RL. 2005. The composition of *Dalea formosa* oil determined by steam distillation and by solid phase microextraction. *J Essen Oil Res* **17**: 645–647.

Lucero ME, Fredrickson EL, Estell RE, Morrison AA and Richman DB. 2006. Volatile composition of *Gutierrezia sarothrae* (broom snakeweed) as determined by steam distillation and solid phase microextraction. *J Essen Oil Res* **18**: 121–125.

Lucero ME, Estell RE and Fredrickson EL. 2009. Composition of *Ceanothus gregii* oil as determined by steam distillation and solid-phase microextraction. *J Essen Oil Res*. (in press)

Medina AL, Lucero ME, Holguin FO, Estell RE, Posakony JJ, Simon J and O'Connell MA. 2005. Composition and Antimicrobial Activity of *Anemopsis californica* Leaf Oil. *J Agr Food Chem* **53**: 8694–8698.

NIST Mass Spec Data Center SES. 2005a. *Mass Spectra*, 6th edn. National Institute of Standards and Technology: Gaithersburg MD.

NIST Mass Spec Data Center SES. 2005b. *Retention Indices*, 6th edn. National Institute of Standards and Technology: Gaithersburg MD.

Tellez MR, Estell RE, Fredrickson EL and Havstad KM. 1997a. Essential oil of *Flourensia cernua* DC. *J Essen Oil Res* **9**: 619–624.

Tellez MR, Estell RE, Fredrickson EL and Havstad KM. 1997b. Essential oil of *Dyssodia acerosa* DC. *J Agri Food Chem* **45**: 3276–3278.

Tellez MR, Estell RE, Fredrickson EL and Havstad KM. 1998. Essential oil of *Chrysothamnus pulchellus* (Gray) Greene ssp. *pulchellus*. *J Essen Oil Res* **10**: 201–204.